Study of Protein Fluctuation with an Effective Inter-C^{α} Atomic Potential Derived from Average Distances between Amino Acids in Proteins

TAKESHI KIKUCHI

Frontier Research Center for Computational Sciences, Science University of Tokyo, 2641 Yamazaki, Noda, Chiba 278-8510, Japan

Received 3 March 1998; accepted 8 January 1999

ABSTRACT: The conformational dynamics around the native structure of bovine pancreatic trypsin inhibitor (BPTI) in both the oxidized and reduced forms was analyzed by a Monte Carlo method using an approximate residue–residue potential derived from the statistics of average distances between C^{α} atoms of residues as proposed by the present author (T. Kikuchi, *J Comput Chem* 1996, 17, 226–237). The results from using this effective potential are similar to those from molecular dynamics simulations, taking all atoms into account, and are consistent with temperature factors from an X-ray analysis and disulfide formation from a kinetic experiment. This agreement suggests that the essential nature of the potential energy surface formed by the potential around the native structure closely mimics the actual energy landscape within the resolution of C^{α} atomic fluctuation. Furthermore, it is expected that the potential we found can describe the basic properties of folding kinetics. Examination of the fluctuation property of the native structure of BPTI threaded by a sequence from cytochrome b562 reveals differences specific to the sequence and this result also shows that the dynamical properties obtained in our calculations are not only ascribed to the geometrical constraints of the initial conformation but also the force field specifically produced by a sequence. J Comput Chem 20: 713–719, 1999

Keywords: interresidue effective potential; Monte Carlo simulations; protein dynamic; bovine pancreatic trypsin inhibitor; molten globule

Correspondence to: T. Kikuchi; e-mail: kikuctal@rs.noda.sut.ac.jp

 CCC 0192-8651 / 99 / 070713-07

Introduction

Many theoretical studies on protein structures have had, as their goals: (1) prediction of protein tertiary structures from only their sequences; and (2) elucidation of how a protein folds into its native structure taking a biological not an astronomical time (i.e., to solve Levinthal's paradox[1]). These two problems are, of course, related. Both knowledge-based interresidue potentials (potentials of mean force[2–9]) and all-atom potentials[10,11] have been used in solving these problem. In particular, potentials of mean force are frequently applied to threading of a sequence to template structures, that is, the so-called 3D–1D methods, as a fitness function to recognize the native fold of a protein.[2–5,12,13] These methods worked successfully for several problems,[14,15] but their success was limited by several factors; for example, alignment accuracy between a sequence and a 3D structure.[16] Improvements of potential functions are also required to detect detailed packing interactions of residues. For the latter problem, similar potentials or more simplified residue–residue interaction potentials are used to study behavior of coarse-grained lattice proteins in order to study kinetic properties of protein folding.[17–26] By virtue of the simplicity of the models, these studies may suggest that the native structure of a protein has enough stability and kinetic accessibility on its energy landscape for folding within reasonable time.[22–24] Such simplified potentials are often used to understand the more detailed dynamical features of protein folding.[25,26] The fluctuations of the proteins around their native structures have already been studied in detail using simplified potentials,[27–29] and they showed that the simulated dynamics agreed with the data of X-ray temperature factors as far as the dynamical properties of the near-native structures were concerned. However, it is not clear to what extent such an approximate potential can describe real dynamics of proteins beyond the local fluctuation around the native structures. One could ask if there are proper artifacts associated with potentials derived from statistics of protein structures and of lattice models in dynamics of proteins.

Recently, we proposed a new interresidue potential based on average distances between C^{α} atoms of amino acids in proteins with known structure and attempted to sample structures with this potential using the sequence of BPTI.[9] It was found that these sampled structures could be categorized into a few classes, one of which contains structures similar to the native structure of BPTI, where, by ''similar,'' we mean about 8-Å root-mean-square deviation (rmsd). This finding demonstrates that this potential possesses an energy landscape that leads to a class of structures similar to the native structure of a protein, although the resolution is somewhat low. The present potential cannot be used to predict tertiary structures of proteins precisely with sufficient accuracy, for example a level of 4-Å rmsd because of this nature of the low resolution. However, our potential has the advantage that it is defined as an analytical function with continuous coordinates (off-lattice model). Therefore, a wide range of dynamics of protein conformations is also easily treated with this potential, in principle. Thus, it is of interest to analyze the coarse-grained dynamical feature of conformational transition of a protein including the structural transition to a molten state by use of our effective potential. In this study, we analyze conformational fluctuation of the native structure of a well-known protein, bovine pancreatic trypsin inhibitor (BPTI), with our effective potential.[9] We examine whether our potential correctly describes the fluctuation dynamics of oxidized (disulfide crosslinked in the native form) BPTI around the native structure by comparison with the X-ray temperature factor data and results of molecular dynamics (MD). Then, we make further comparison of the present simulations of the reduced (no disulfide constraints) BPTI with the MD calculations with all atoms. In particular, we compared our results of simulation for the reduced form with the MD results performed by Daggett and Levitt[30] as a reference point of the actual dynamics of BPTI. The fluctuation properties of a protein conformation around its native structure might be ascribed to its geometrical constraints. Therefore, we also performed the same simulations with a sequence completely different from that of BPTI and compared the findings with the BPTI results. This problem might relate to the compatibility of the dynamical mode of a sequence with a geometrical constraint.

Method

The same method and models in ref. 9 were used in the present work; that is, the single Gauss-

ian potential[9] was applied to BPTI. Accordingly, two successive C^{α} atoms were connected by a virtual bond with a length of 3.8 Å, and the effect of a residue was effectively included in the interactions between C^{α} atoms. A conformation of this model protein can be described by a set of bond angles and dihedral angles. The native disulfide bonds were incorporated in the model in the same way as in the previous work.[9]

The procedure for simulation is as follows. The standard Metropolis Monte Carlo (MC) method[31] was employed. A bond angle and a dihedral angle were randomly selected during the MC simulation by the following amounts.

$$\text{Bond angle} = \pi \times \eta \times \delta_1$$
$$\text{Dihedral angle} = 2\pi \times \eta \times \delta_2$$

Here, η stands for a random number within the interval $[0, 1]$ and δ_1 and δ were determined to obtain near-equilibrium fluctuation of energy around the relaxed structure (see below) during the simulation time. In the present case, this step was iterated 40,000 times taking 0.03 as the value for δ_1 and δ_2. (The acceptance ratio was about 30% to 50% during a whole Monte Carlo procedure.) An rmsd of each C^{α} atom from the relaxed structure during the simulation was monitored and we confirmed that sufficient fluctuation of a conformation was obtained in the simulation as an equilibrium state; that is, a conformation did not fall into a frozen state. (We normally obtained 0.5−1.5-Å-rms fluctuation for each residue at kT = 0.6 in equilibrium states.) We performed the calculations for BPTI with and without disulfide bonds (i.e., oxidized and reduced forms) for several values of temperature parameter kT. In reduced form, the disulfide constraints were not implemented.

A structure from the X-ray structure of BPTI relaxed by our potential was used as the initial structure for all present simulations. Actually, if we adopt the X-ray structure itself as an initial conformation for the simulations, this structure might be strained within a narrow and deep well on a coarse-grained potential and only small local fluctuation around the X-ray structure would be observed. Therefore, we consider the relaxation of the X-ray conformation to be a prerequisite, especially when we use an approximate potential. For relaxation, 40,000 MC steps were performed starting from the X-ray structure at kT = 1.2. The lowest energy structure was selected from these steps, followed by further energy minimization with the present potential. This structure (Fig. 1) was defined as the relaxed structure. The rmsd was 2.27 Å, as compared with the C^{α} of the X-ray structure.

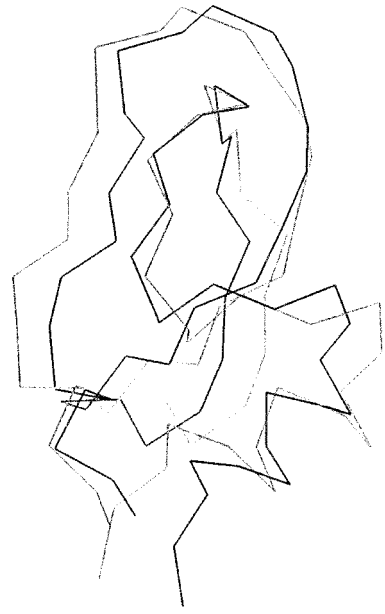

FIGURE 1. The relaxed structure of BPTI using the present potential (gray line) compared with the X-ray structure (black line). The rms deviation of C^{α} atoms is 2.27 Å.

Results

Table I summarizes the values of radii of gyration (R_g) of conformations of BPTI of the final structures after the MC simulations of oxidized and reduced forms with T = 0.6 and 1.2. (For the oxidized form, T = 1.2 and T = 0.6 gave essentially the same result.) For comparison, we also present the results of the MD simulation for BPTI, including all atoms with water performed by Dagett and Levitt[30] (DL) at T = 298, 423, and 429 K. The R_g value calculated with the present potential for oxidized BPTI at kT = 0.6 is 11.1, which is close to the values of 11.6 and 11.7 obtained by DL at T = 298 and 423. R_g values for reduced BPTI are 11.7 and 12.2 at kT = 0.6 and 1.2, respectively. These values are again close to those by DL at T = 298, 423, and 429 (i.e., 11.9, 12.1, and 12.6). The DL results indicate that, compared with the oxidized form, the average volume of the reduced BPTI was increased by 11% to 25%, and this increase compares favorably with the experimentally determined increase for the molten globule state (10% to 25%).[32−34] According to DL, the increase

TABLE I.
Comparison of Properties of BPTI Calculated in the Present Work and Results of Molecular Dynamics Study.[30]

	Present study			MD[30]				
	Oxidized, kT	Reduced (kT)		Oxidized (T)		Reduced (T)		
	0.6 kT	0.6	1.2	298 K	423 K	298 K	423 K	498 K
Radius of gyration, R_g (Å)	11.1	11.7	12.2	11.6	11.7	11.9	12.1	12.6
rms from X-ray structure (Å)	2.4	3.4	4.7	1.7	2.7	2.5	3.9	5.1

in volume is not caused by solvent penetration, but by conformational adjustments. It is noteworthy that the increase in volume of the reduced form of BPTI is essentially reproducible by our system.

The overall fluctuation of BPTI can be described by the C^α rmsd values from the relaxed structure of BPTI averaged over a whole simulation. Figure 2 shows the C^α rmsd after the simulation of the oxidized form at kT = 0.6. The data were obtained from *in vacuo* MD at T = 298 K,[35] the normal mode analysis[36]; the MD simulation with water at T = 298K[30] and the temperature factor of the X-ray analysis are also illustrated in Figure 2. The low fluctuating regions in BPTI correspond to the regular (secondary) structures; that is, 2–7 and 47–56 of α-helices and 18–24 and 29–35 of β-sheets. On the other hand, the highly mobile parts are loop regions between these regular structures. These properties of fluctuation dynamics coincide with other simulations and the experiment. In particular, our simulation with the present protocol for the oxidized form shows good agreement with the results of in vacuo MD simulation[35] and normal mode analysis.[36] A similar trend also appears in the MD simulation of reduced BPTI in the DL results.[30] In Figure 3, we present the C^α rmsd values of reduced BPTI from the relaxed structure at kT = 0.6 and 1.2 with the results of the DL simulations at T = 423 K and 498 K. Again, the low fluctuating regions in reduced BPTI correspond to the regular structures, and our simulation at kT = 1.2 agree well with the DL simulations (especially at T = 423 K). Furthermore, the results of our calculations and the DL simulations both demonstrate that the properties of the highly flexible region are not caused by the disulfide constraint only.

The high mobility of the N- and C-terminal regions, 1–13 and 47–50, at kT = 0.6 (Fig. 3), im-

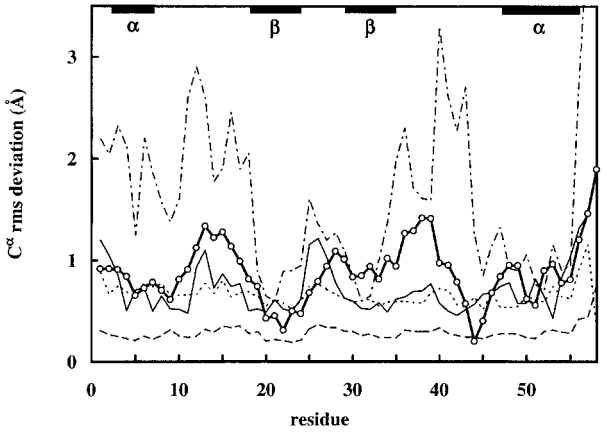

FIGURE 2. C^α rms deviation from the relaxed structure of oxidized BPTI over the whole simulation at kT = 0.6 (line with open circles), the results of the MD simulations *in vacuo* (solid line) at T = 298 K,[35] with water (dotted / dashed line) at T = 298 K,[30] normal mode analysis[36] (dotted line), and the temperature factor in the X-ray structure (dashed line). The filled bars denote the location of secondary structures (i.e., α-helices or β-strands).

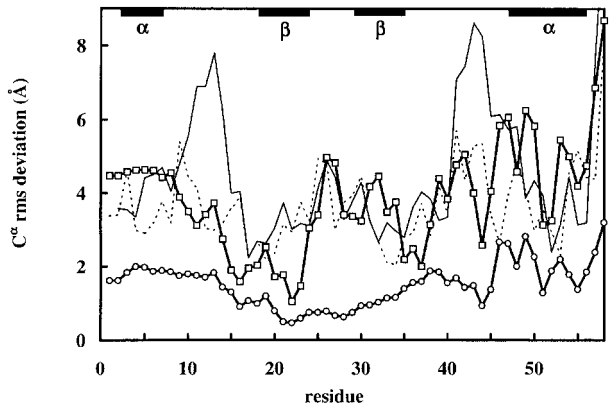

FIGURE 3. C^α rms deviations from the relaxed structure of reduced BPTI over the whole simulation at kT = 0.6 (line with open circles) and 1.2 (line with open squares) and results of the MD simulation with water at T = 423 K (dotted line) and T = 498 K (solid line).[30]

plies a higher degree of freedom of both terminals in the reduced form. The trend is essentially similar in the simulation at kT = 1.2, but the mobility of both ends increases remarkably and the motion around residue 30[37] becomes high, probably due to the reduction of disulfide 30–51. Similar effects also appear in the simulations in DL; that is, the motion around the N-terminal residues 10–15 and around the C-terminal residues 45–50 is high, and the remarkable peak around Cys 30 emerges at T = 423 K and 498 K.

Figure 4 shows the superposition of the final structure of the simulation of reduced BPTI at kT = 1.2 with the X-ray structure. It can be seen from this figure that the N- and C-terminal portions have moved significantly from the X-ray structure. This is also observed in the simulation done by DL (see Fig. 1 in ref. 30).

The fluctuation profile of the conformation of BPTI might be intrinsic to the tertiary structure itself; that is, the fluctuation around its native structure might not be specific to its sequence and might be insensitive to details of the model potential. To examine this matter, we performed the same simulation using another sequence with no relation to BPTI. The sequence of cytochrome b562 from 1 to 58 was employed in the present study because the tertiary structure of cytochrome b562 does not have any similarity to that of BPTI. Cytochrome b562 is a pure α-protein consisting of 106 amino acids, whereas BPTI is a small ($\alpha + \beta$)-protein. The sequence of cytochrome b562 1–58 was threaded into the tertiary structure of BPTI, but the residues at the same positions corresponding to Cys's forming the native disulfide bridges in BPTI were replaced by Cys's in the sequence of cytochrome b562. In this case, our purpose was to examine whether the fluctuation profile depends not only on the geometric constraints but also on the force field intrinsic to the sequence. Therefore, we did not perform the conformational relaxation of the X-ray structure of cytochrome b562. Figure 5 represents the C^{α} rmsd values of the reduced form of this fictitous protein and reduced BPTI at kT = 0.6. The overall fluctuation of the protein with the sequence from cytochrome b562 is larger than that of BPTI. Compared with BPTI, which shows high mobility of both terminals, the plot for cytochrome b562 reveals large fluctuation of the central part of the protein rather as opposed to the N- and C-terminals. Furthermore, the relatively low mobile regions in BPTI exhibit larger fluctuation in cytochrome b562, especially β-strand 18–24. On the other hand, the region with high mobility in BPTI (45-50) fluctuates little in this fictitious protein. In summary, the conformation of reduced BPTI changes via large fluctuation of both termini, but the structure threaded by the sequence from cytochrome b562 seems to break gradually at the central part of the protein. From these results, we conclude that the fluctuation profiles in Figure 5

FIGURE 4. The superposition of the final structure of reduced BPTI (gray line) calculated with the C^{α} potential at kT = 1.2 with the X-ray structure (black line). The deviation of the N-terminal part is noteworthy. The C-terminal is also fairly different from the X-ray structure.

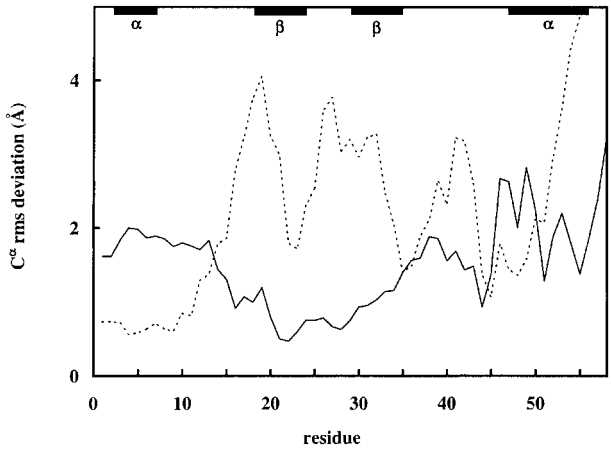

FIGURE 5. C^{α} rms deviations from the relaxed structure of reduced BPTI over the whole simulation at kT = 0.6 (solid line), and the same plot of the tertiary structure of reduced BPTI threaded by the sequence from cytochrome b562 1–58 (dotted line).

reflect the difference between the shapes of the potential energy surfaces produced by the two different sequences.

To relate our simulations to the unfolding (or refolding) process of BPTI, we analyzed changes of the distances between the Cys pairs in simulations. Table II presents the distances between the Cys pairs of the disulfides in the native form after the simulations of oxidized and reduced BPTI as a function of temperature. Increasing the kT value, the distance between the Cys pairs of the disulfide bridges (i.e., 5–55 and 14–38) becomes larger in the reduced form (up to 17.04 and 10.57 at kT = 1.2, respectively). On the other hand, the distance of Cys pair 30–51 maintains its short distance (e.g., 7.33 at kT = 1.2), which is relatively close to the oxidized distance of 5.52. Cys-30 and Cys-51 are close to each other during the simulation, suggesting that this Cys pair could easily form a disulfide in the early stage of the refolding process. This result corresponds to the experimental results of refolding kinetics of BPTI.[38,39]

Discussion

It is hoped that fluctuation of protein conformation can be described by an analytical function with a continuous coordinate system like the present potential. Many of the potentials proposed so far[6–8,17–26] used lattice models and/or discrete functions of potentials. Recently, the continuous potential function was also applied to the calculations of the dynamics around the native structures of the proteins and the results show good agreement with the experiments.[27–29] In this article, we demonstrated that the present potential can reproduce the properties of the conformational fluctuation of BPTI around the native structure in addition to further away from the native structure, as shown in Table I and Figures 2 and 3. In particular, the point of interest is that our results are consistent with the MD simulations of reduced BPTI by DL (Fig. 3). These results suggest that the present potential gives an energy landscape similar to that of the all-atom interatomic potential energy function. This is true for both oxidized and reduced BPTI within the extent of the resolution of C^{α} motion. Furthermore, DL could simulate the process of the expansion of BPTI by 10% to 25% by about 500-ps simulations.[30] Likewise, we could reproduce the similar process with our potential as demonstrated in Table I. Thus, our simulation with the present potential of 40,000 MC steps can approximate such dynamical processes of BPTI in around several 100 ps.

Our results for temperature dependence of the distances between disulfide-forming Cys pairs in reduced BPTI (Table II) also correspond qualitatively to the results of kinetics of the disulfide bond formation. Our simulations pertain to only the structural feasibility of Cys pairs and do not include the actual dynamics of disulfide formation. However, our results demonstrate that Cys pair 30–51, which forms first in the refolding kinetics of BPTI, as shown in the Creighton experiment,[38] exists separated by a relatively short distance compared with the other Cys pairs (see Table II). This suggests that Cys pair 30–51 forms a disulfide bond more easily than others. Hence, it is expected that our potential will be able to simulate the dynamics of disulfide bond formation in the refolding process of BPTI, although it will require a much longer simulation time and careful analyses. We are currently preparing such analyses.

We also demonstrated that the fluctuation profile of reduced BPTI is not only attributed to the tertiary structure itself but also depends on the amino acid sequence. That is, the potential energy surface around the native structure produced by the sequence of BPTI shows a rather different shape compared with that produced by the sequence of cytochrome b562. This effect might be remarkable in the case of large fluctuation such as in the reduced form of BPTI. Furthermore, it is assumed that, if a sequence is not compatible with

TABLE II.
Distance (Å) between Cysteines Forming Native Disulfides in BPTI.

Residue number of Cys		Reduced (kT)			
Cys (first)	Cys (second)	0.6	0.9	1.2	Oxidized, 0.6 (kT)
5	55	10.79	12.16	17.04	5.58
14	38	6.98	7.94	10.57	5.54
30	51	5.73	6.43	7.33	5.52

the proper tertiary structure, the mobility of several parts of this protein is largely different from proper mobility. That is, it might suggest that dynamical compatibility should also be taken into consideration in the prediction of sequence-tertiary structure compatibility by threading.

We conclude that the energy landscape of BPTI formed by the present potential can describe the dynamical behavior around its native structure correctly within the resolution of the C^{α} atom fluctuation. In this study, we used the fluctuation of each residue in comparisons of the two simulations. However, to study properties of collective motions of a protein, it might be better if fluctuation are decomposed into modes and the modes be used for comparisons of dynamics of protein conformational fluctuation.[40] It is also assumed that the present potential can predict correctly more drastic dynamics of a protein, like transition between native and partially disordered states. For example, the results of the temperature dependence of the distances of Cys pairs suggest that our potential can be used for the prediction of kinetic property of protein unfolding. The next step of the present work will be analysis of such processes using our potential and comparing the results with experiment. However, in the cooperative transition between the native and molten globule states, fixation of the side-chain conformations might be essential, as several investigators have noted.[41–43] Therefore, to describe this transition correctly, we must incorporate the effect of side-chain motion explicitly. We are currently in the process of experimentation in this area.

References

1. Levinthal, C. J Chim Phys 1968, 65, 44.
2. Sippl, M. J. Curr Opin Struct Biol 1995, 5, 229.
3. Wodak, S. J.; Rooman, M. J. Curr Opin Struct Biol 1993, 3, 247.
4. Matsuo, Y.; Nishikawa, K. Proteins 1995, 23, 370.
5. Bryant, S. H.; Larence, C. E. Proteins 1993, 16, 92.
6. Sippl, M. J. J Mol Biol 1990, 213, 859.
7. Wilson, C.; Doniach, S. Proteins 1989, 6, 193.
8. Miyazawa, S.; Jernigan, R. L. Macromolecules 1985, 18, 534.
9. Kikuchi, T. J Comput Chem 1996, 17, 226.
10. Vásques, M.; Némethy, G.; Scheraga, H. A. Chem Rev 1994, 94, 2183.
11. Hansmann, U. H. E.; Okamoto, Y. Chem Phys Lett 1996, 259, 321.
12. Bryant, S. H.; Altschul, S. F. Curr Opin Struct Biol 1995, 5, 236.
13. Bowie, J. U.; Eisenberg, D. Curr Opin Struct Biol 1993, 3, 437.
14. Moult, J.; Pedersen, J. T.; Judson, F.; Fidelis, K. Proteins 1995, 23, ii.
15. Lemer, C. M.-R.; Rooman, M. J.; Wodak, S. J. Proteins 1995, 23, 337.
16. Bryant, S. H. Protein 1996, 26, 172.
17. Šali, A.; Shakhnovich, E.; Karplus, M. Nature 1994, 369, 248.
18. Chan, H. S.; Dill, K. A. J Chem Phys 1994, 100, 9238.
19. Shakhnovich, E. I. Phys Rev E 1994, 72, 3907.
20. Hao, M.-H.; Scheraga, H. A. J Phys Chem 1994, 98, 4940.
21. Socci, N. D.; Onuchic, J. N. J Phys Chem 1994, 101, 1519.
22. Onuchic, J. N.; Wolynes, P. G.; Luthey-Schulten, Z.; Socci, N. D. Proc Natl Acad Sci USA 1995, 92, 3626.
23. Socci, N. D.; Onuchic, J. N.; Wolynes, P. G. J Chem Phys 1996, 104, 5860.
24. Camacho, C. J.; Thirumalai, D. Proc Natl Acad Sci USA 1993, 90, 6369.
25. Covell, D. G. J Mol Biol 1994, 235, 1032.
26. Šali, A.; Shakhnovich, E. I.; Karplus, M. J Mol Biol 1994, 235, 1614.
27. Haliloglu, T.; Bahar, I. Proteins 1998, 31, 271.
28. Bahar, I.; Erman, B.; Haliloglu, T.; Jernigan, R. L. Biochemistry 1997, 36, 13512.
29. Bahar, I.; Wallqvist, A.; Covell, D. G.; Jernigan, R. L. Biochemistry 1998, 37, 1067.
30. Daggett, V.; Levitt, M. Proc Natl Acad Sci USA 1992, 89, 5142.
31. Metropolis, N.; Rosenbluth, A.; Rosenbluth, M.; Teller, A.; Teller, E. J Chem Phys 1953, 21, 1087.
32. Kuwajima, K. Proteins 1989, 6, 87.
33. Ptitsyn, O. B. J Protein Chem 1987, 6, 273.
34. Amir, D.; Haas, F. Biochemistry 1988, 27, 8889.
35. MacCammon, J. A.; Gein, B. R.; Karplus, M. Nature 1977, 267, 585.
36. Go, N.; Noguchi, T.; Nishikawa, T. Proc Natl Acad Sci USA 1983, 80, 3696.
37. In this case, the peak appears at the residue 32.
38. Creighton, T. E. Prog Biophys Mol Biol 1978, 33, 231.
39. Gilbert, H. F. In: Pain, R. H., ed. Mechanisms of Protein Folding; Oxford University Press: Oxford, 1994; p 104.
40. Higo, J.; Umeyama, H. Protein Eng 1997, S10, 373.
41. Kolinski, A.; Galazka, W.; Skolnick, J. Proteins 1996, 26, 271.
42. Shakhnovich, E. I.; Finkelstein, A. V. Biopolymers 1989, 28, 1667.
43. Finkelstein, A. V.; Shakhnovich, E. I. Biopolymers 1989, 28, 1681.